Machine Learning and Data Science

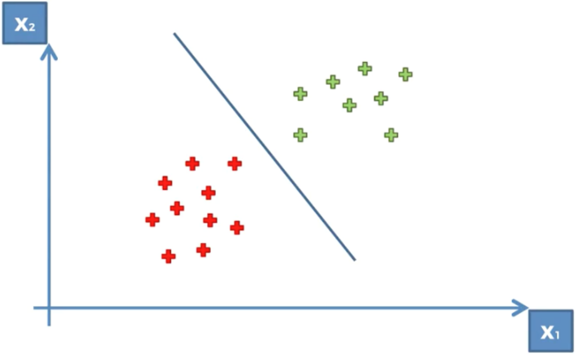
Machine à vecteurs de support, astuce du noyau (kernel SVM)

Bassem Ben Hamed

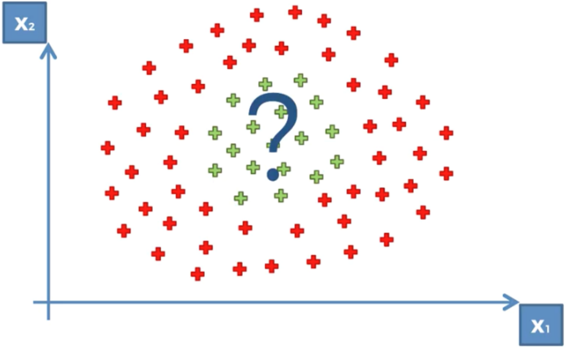
Juillet 2018

Kernel SVM Intuition

SVM sépare bien ces points



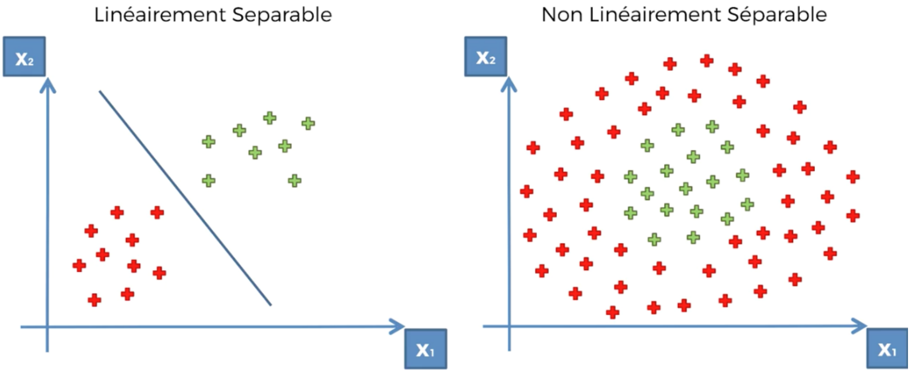
Et ces points ?



Pourquoi ?

Parce que les points d'observation ne sont pas LINÉAIREMENT SEPARABLES

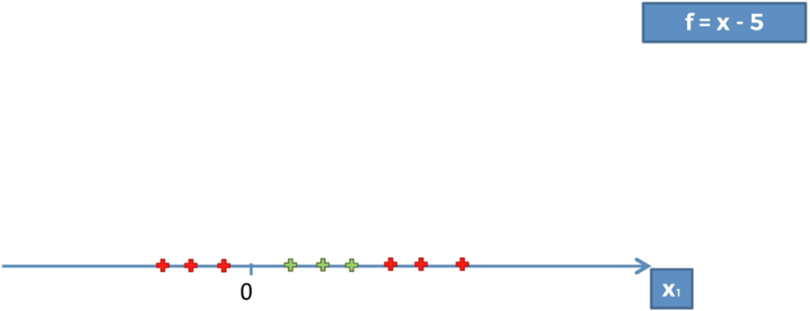
Linéairement Séparable



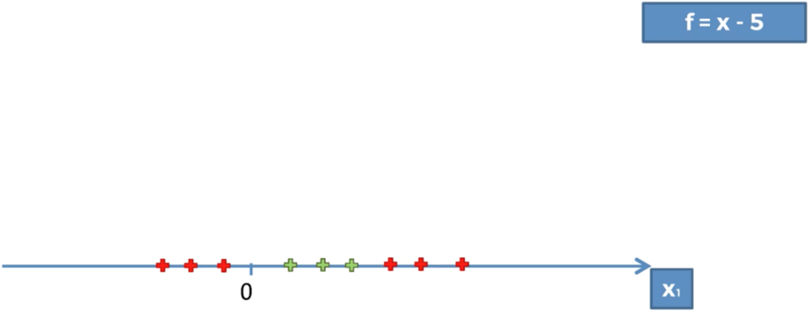
Espace de plus grande dimension

Mapping vers une plus grande dimension

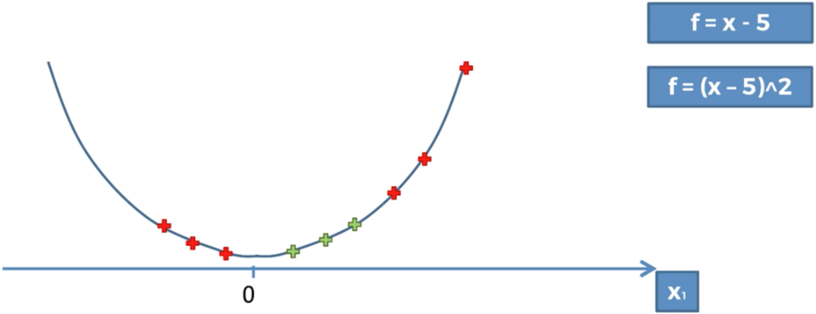
Mapping vers une plus grande dimension



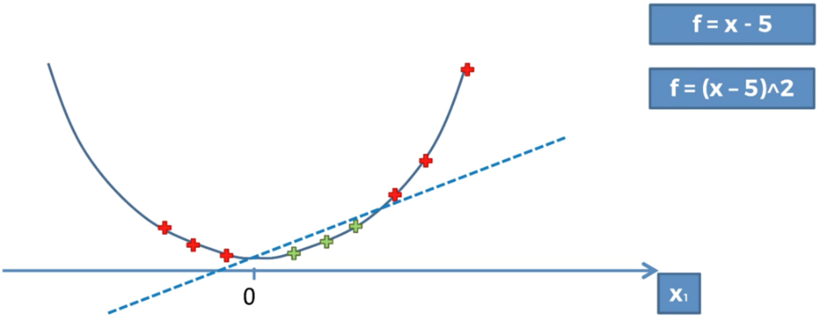
Mapping vers une plus grande dimension



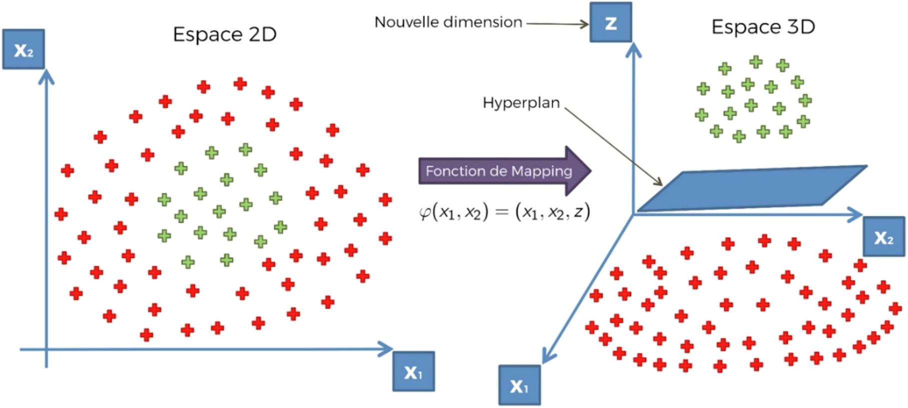
Mapping vers une plus grande dimension



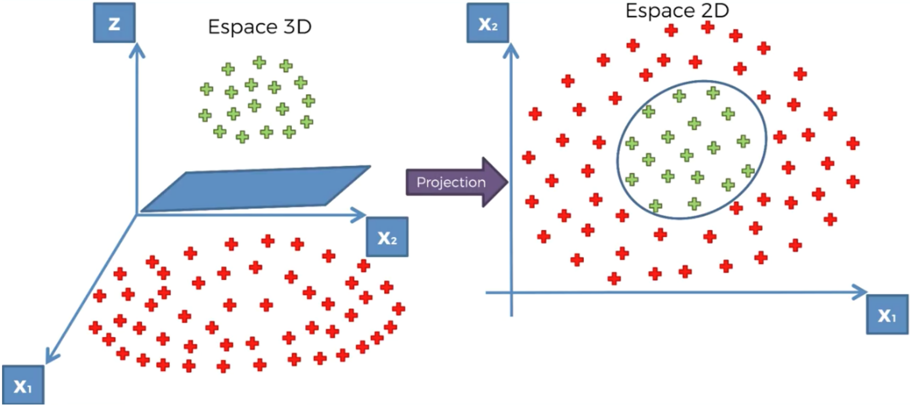
Mapping vers une plus grande dimension



Mapping vers une plus grande dimension



Re-Projection vers Espace 2D

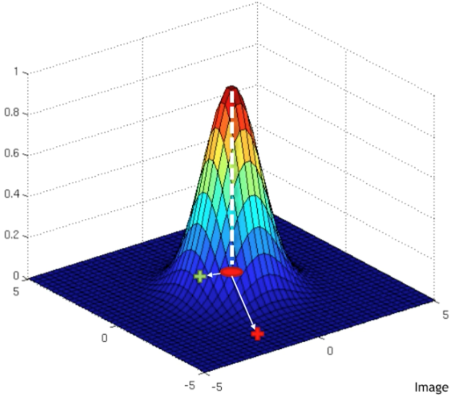


Mais il y a un petit problème...

Le Mapping vers un espace de plus grande dimension peut demander trop de calculs

La solution Kernel

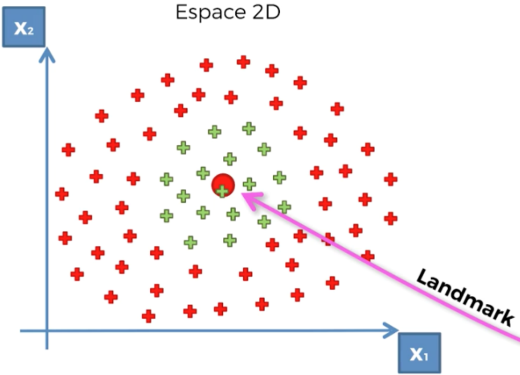
Gaussian RBF Kernel

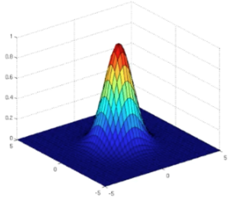


$$K(\vec{x}, \vec{l}^i) = e^{-\frac{\left\|\vec{x}-\vec{l}^i\right\|^2}{2\sigma^2}}$$

Image source: http://www.cs.toronto.edu/~duvenaud/cookbook/index.html

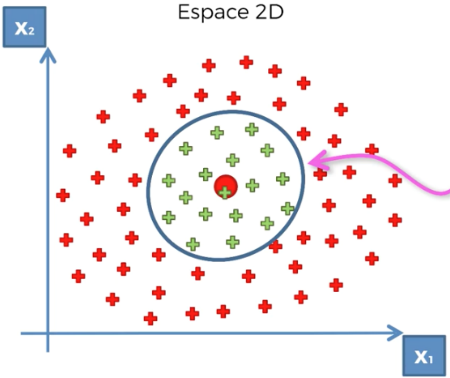
Gaussian RBF Kernel

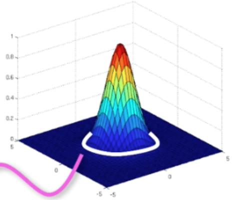




$$K(\vec{x},\vec{l}^{i}) = e^{-\frac{\left\|\vec{x}-\vec{l}^{i}\right\|^{2}}{2\sigma^{2}}}$$

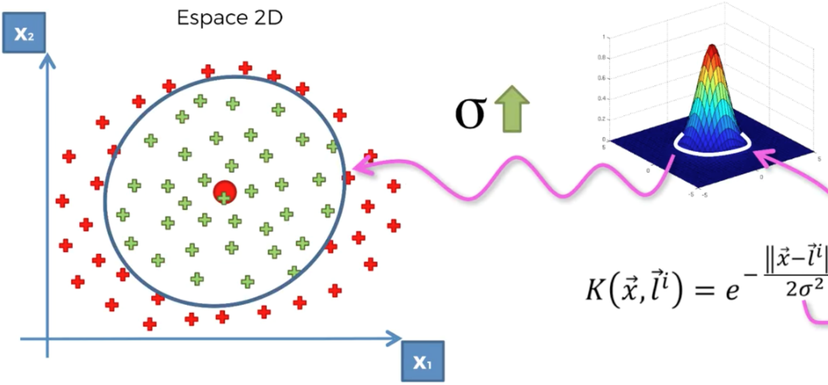
Gaussian RBF Kernel





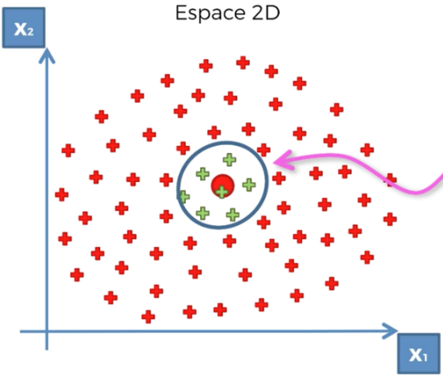
$$K(\vec{x}, \vec{l}^i) = e^{-\frac{\|\vec{x}-\vec{l}^i\|^2}{2\sigma^2}}$$

Gaussian RBF Kernel

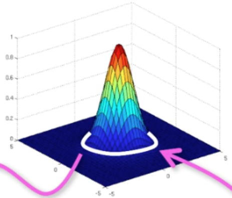


$$K(\vec{x}, \vec{l}^i) = e^{-\frac{\|\vec{x}-\vec{l}^i\|^2}{2\sigma^2}}$$

Gaussian RBF Kernel

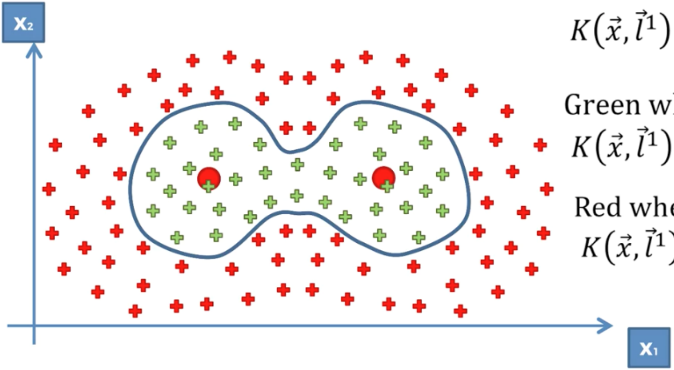


$\sigma \downarrow$



$$K(\vec{x}, \vec{l}^i) = e^{-\frac{\left\|\vec{x}-\vec{l}^i\right\|^2}{2\sigma^2}}$$

Gaussian RBF Kernel



$$K(\vec{x}, \vec{l}^1) + K(\vec{x}, \vec{l}^2)$$

(Simplified Formula)

Green when:

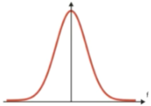
$$K(\vec{x}, \vec{l}^1) + K(\vec{x}, \vec{l}^2) > 0$$

Red when:

$$K(\vec{x}, \vec{l}^1) + K(\vec{x}, \vec{l}^2) = 0$$

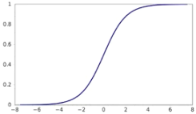
D'autres Kernels

D'autres Kernels



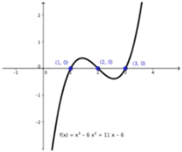
Gaussian RBF Kernel

$$K\left(\vec{x}, \vec{l}^{i}\right) = e^{-\frac{\left\|\vec{x}-\vec{l}^{i}\right\|^{2}}{2\sigma^{2}}}$$



Sigmoid Kernel

$$K(X,Y) = \tanh(\gamma \cdot X^{T}Y + r)$$



Polynomial Kernel

$$K(X,Y) = (\gamma \cdot X^{T}Y + r)^{d}, \gamma > 0$$